

WHAT IS CLAIMED:

1 1. A method for reducing memory latency in a multi-node architecture, comprising:
2 receiving a speculative memory read request at a home node before results of a
3 cache coherence protocol are determined; and
4 initiating a read to memory to complete the speculative memory read request.

1 2. The method of claim 2, further comprising:
2 buffering results of the read to memory.

1 3. The method of claim 2, further comprising:
2 dropping the results of the read to memory on a buffer full condition or if a cancel
3 command is received.

1 4. The method of claim 3, further comprising:
2 if a confirm command is received after results of the read to memory are dropped,
3 initiating a second read to memory to complete a memory read request.

1 5. The method of claim 4, further comprising:
2 forwarding results of the second read to memory to a requester.

1 6. The method of claim 3, further comprising:
2 if a confirm command is received before results of the speculative read are
3 dropped, forwarding the results of the read to memory to a requester.

1 7. The method of claim 6, wherein the speculative memory read request is issued by
2 the requesting node.

1 8. The method of claim 6, further comprising:
2 receiving the results of the read to memory at the coherence agent; and
3 forwarding the results of the read to memory to the requesting node.

1 9. A method for reducing memory latency, comprising:
2 issuing a speculative memory read request to a home node before results of a
3 cache coherence protocol are determined; and
4 initiating the cache coherence protocol.

1 10. The method of claim 9, further comprising:
2 updating a memory status relating to the results in a table after the results of the
3 cache coherence protocol are determined.

1 11. The method of claim 9, wherein initiating the cache coherence protocol
2 comprising:
3 initiating a status look-up to determine the caching status of the requested
4 memory.

1 12. The method of claim 11, further comprising:
2 issuing a confirm command to the home node if the caching status is determined
3 to be in an invalid state or shared state.

1 13. The method of claim 11, further comprising:
2 snooping a node with the exclusive copy of the requested memory cached.

1 14. The method of claim 13, further comprising :
2 determining whether the exclusive copy of the requested memory is clean or dirty.

1 15. The method of claim 14, further comprising:
2 issuing a confirm command to the home node if the exclusive copy of the
3 requested memory is clean.

1 16. The method of claim 14, further comprising:
2 issuing a cancel command to the home node if the exclusive copy of the requested
3 memory is dirty.

1 17. The method of claim 13, further comprising:
2 receiving a snoop result, wherein the snoop result includes a copy of the requested
3 memory; and
4 updating a memory status relating to the requested memory in a table.

1 18. The method of claim 17, further comprising:
2 receiving the requested memory; and
3 forwarding the requested memory to a requesting node.

1 19. A home node for responding to read requests in a multi-node architecture
2 including a plurality of nodes, the home node comprising:
3 a processor;

4 a memory; and
5 a node controller coupled to the processor and memory, the node controller
6 adapted to:
7 receive a speculative memory read request from a requester in the multi-
8 node architecture before a cache coherence protocol is resolved, and
9 initiate a read to memory to complete the speculative memory read
10 request.

1 20. The home node of claim 19 further comprising:
2 a buffer adapted to buffer the results of the read to memory.

1 21. The home node of claim 20, wherein the results of the read from memory are
2 dropped from the buffer on a buffer full condition or upon receiving a cancel command.

1 22. The home node of claim 20, wherein the node controller responsive to a confirm
2 is adapted to forward the results of the read to memory to the requester.

1 23. The home node of claim 20, wherein the node controller responsive to a cancel
2 command is adapted to drop the data specified by the speculative read command.

1 24. A system comprising:
2 a node including a node controller adapted to control a plurality of processors
3 resident in the node, wherein the node controller adapted to receive a speculative read
4 request before results of a coherence protocol are determined and the node controller
5 adapted to read data specified by the speculative read command from memory; and

6 a coherence agent coupled to the at least one node, the coherence agent including
7 a coherence controller adapted to determine the results of the coherence protocol and
8 adapted to forward a cancel command or a confirm command to the node after the results
9 of the coherence protocol are determined.

1 25. The system of claim 24, wherein the node controller responsive to the confirm
2 command issued by the coherence controller is adapted to send the data read from
3 memory to the coherence controller.

1 26. The system of claim 24, wherein the node controller responsive to the cancel
2 command issued by the coherence controller is adapted to drop the data read from
3 memory.

1 27. The system of claim 24, further comprising:
2 a requesting node adapted to send a data read request to request data identified by
3 a memory address included in the data read request.

1 28. The system of claim 27, wherein the speculative read request is sent by the
2 requesting node.

1 29. The system of claim 24, wherein the speculative read request is sent by the
2 switching agent.